SYSTEMS OPTIMIZATION LABORATORY
DEPARTMENT OF OPERATIONS RESEARCH
STANFORD UNIVERSITY
STANFORD, CALIFORNIA 94305-4022

A Build-Up Interior Method for Linear Programming: Affine Scaling Form

by

George B. Dantzig and Yinyu Ye

TECHNICAL REPORT SOL 90-4

February 1990

Research and reproduction of this report were partially supported by the National Science Foundation grants DDM-8814253, DMS-8913089; Office of Naval Research grant N00014-89-J-1659 and the Department of Energy grant DE-FG03-87ER25028.

A BUILD-UP INTERIOR METHOD FOR LINEAR PROGRAMMING: AFFINE SCALING FORM

George B. Dantzig † and *Yinyu Ye* ‡

February 1990

Key words: linear programming, interior method, affine scaling.

Abbreviated title: A Build-Up Interior Method for LP

† Department of Operations Research, Stanford University, Stanford, CA 94305

‡ Department of Management Sciences, The University of Iowa, Iowa City, IA 52242

Abstract

We propose a *build-up* interior method for solving an m equation n variable linear program which has the same convergence properties as their well known analogues in dual affine and projective forms but requires less computational effort. The algorithm has three forms, an *affine scaling* form, a *projective scaling* form, and an *exact* form (that uses pivot steps). In this paper, we present the first of these. It differs from Dikin's algorithm of dual affine form in that the ellipsoid chosen to generate the improving direction $\bar{\Delta}$ in dual space is constructed from only a subset of the dual constraints.

At the start of each major iteration t, we are given an interior iterate y^t. A selection of m dual constraints is made using an "order-columns" rule as to which constraints show "the most promise" of being tight in the optimal dual solution. An ellipsoid centered at y^t is then inscribed in convex region defined by these promising constraints and an improving direction $\bar{\Delta}$ computed that points to the optimal point $y^t + \bar{\Delta}$ on the ellipsoid boundary. Minor cycling within a major iteration is then started.

During a minor cycle, the constraints selected to define the ellipsoid centered at y^t is built up to include the constraint (whenever there is one) that first blocks feasible movement from y^t to $y^t + \bar{\Delta}$. If one blocks, it is used to augment the set of promising constraints and the ellipsoid is revised; the improving direction $\bar{\Delta}$ is recomputed by means of a rank-one update, and the minor cycle repeated until none blocks movement from y^t to $y^t + \bar{\Delta}$. When none blocks, the minor cycling ends. $y^{t+1} = y^t + \bar{\Delta}$ initiates the next major iteration. Major iterations stop when an optimum solution is reached. We prove this will occur in a finite number of iterations.

1. Steps of Dikin's Algorithm

We are concerned with the linear program whose primal form is

$$(\mathcal{P}) \qquad \text{minimize} \quad \bar{z} = cx$$

$$\text{subject to} \quad x \in \{x \in R^n : Ax = b,\ x \geq 0\}, \tag{1.1}$$

where $c \in R^n$, $A \in R^{m \times n}$, and $0 \neq b \in R^m$ are given. Its dual form is

$$(\mathcal{D}) \qquad \text{maximize} \quad \underline{z} = yb$$

$$\text{subject to} \quad y \in \{y \in R^m : c - yA \geq 0\}. \tag{1.2}$$

We denote the dual slack variables by

$$u = c - yA \geq 0. \tag{1.3}$$

When feasible solutions exist for both $(\mathcal{P})$ and $(\mathcal{D})$, then feasible solutions x and (y, u) are optimal for $(\mathcal{P})$ and $(\mathcal{D})$ if and only if the vector

$$\text{Diag}(u)x = 0, \tag{1.4}$$

see Dantzig [4]. Diag(u) denotes the diagonal matrix with diagonal u. We therefore seek feasible solutions for $(\mathcal{P})$ and $(\mathcal{D})$ satisfying complementarity conditions (1.4).

In order to show convergence for our build-up method, we need to refer to specific parts of the proof for Dikin's algorithm, and the best way we found to do so is to first give our own proof of the latter in the next two sections. Proofs of Dikin's affine scaling algorithm [5] were further developed by Adler, Karmarkar, Resende and Veiga [1], Barnes [2], Monma and Morton [8], and Vanderbei, Meketon and Freedman [11], among others. We will present a set of assumptions and inductive conditions upon which a proof of convergence can be based and show how these extend to our variant; hence details of their results with modifications provide a proof that our build-up algorithm also converges to optimal solutions to $(\mathcal{P})$ and $(\mathcal{D})$. We then show that our

starting rules for initiating a major iteration imply convergence in a finite number of iterations.

Each iteration t of the Dikin algorithm starts with an interior dual y^t, solves an ellipsoid subproblem centered at y^t:

$$(\mathcal{E}) \qquad \text{maximize} \quad \underline{z} = yb$$
$$\text{subject to} \quad y \in \mathcal{E} = \{y \in R^m : \|(y - y^t)AD^{-1}\| \leq 1\}, \tag{2.0}$$

where the Euclidean norm of a vector v is denoted by $\|v\|$ and $D = \text{Diag}(u^t)$ denotes the diagonal matrix with diagonal u^t.

The optimal solution to $(\mathcal{E})$ becomes y^{t+1} of the next iteration. We will review later the proof that y^{t+1} lies in the interior of $(\mathcal{D})$. It is computed by (2.2) below. At the same time, a solution x^t is computed by (2.3) which satisfies $Ax^t = b$; x^t may or may not be a feasible solution to $(\mathcal{P})$. As $t \to \infty$, we will show that x^t tends towards being feasible and also becomes more and more complementary to:

$$u^t = c - y^t A > 0. \tag{2.1}$$

The algorithm starts with an interior solution y^1 given. The iterates are then computed by

$$y^{t+1} = y^t + (\Delta b)^{-1/2}\Delta, \tag{2.2}$$

$$x^t = D^{-2}A^T\Delta^T \tag{2.3}$$

where

$$\Delta = b^T(AD^{-2}A^T)^{-1} \quad \text{and} \quad D = \text{Diag}(u^t). \tag{2.4}$$

The superscript ^{T} denotes transpose.

In place of (2.2), an adjusted y^{t+1} is often used that makes a bigger step:

$$\text{adj.} \quad y^{t+1} = y^t + \alpha(\Delta b)^{-1/2}\Delta, \quad \alpha \geq 1,$$

where α is chosen so that adj. y^{t+1} is .9 of the way in the direction Δ from unadjusted y^{t+1} given by (2.2) to the boundary of $(\mathcal{D})$. The proof of convergence is almost exactly the same. For our development, we assume $\alpha = 1$.

2. Motivation for the Build-Up Algorithm

The improved point $y^{t+1} = y^t + \bar{\Delta} = y^t + (\Delta b)^{-1/2}\Delta$ results from solving the ellipsoid constrained subproblem (2.0). The computation of $b^T(AD^{-2}A^T)^{-1}$ involves, however, all the columns of A even though, in practical problems, most columns have little effect on the shape and size of the ellipsoid and hence on the location of the optimal point $y = y^{t+1}$. Still other columns may affect its location but may do so adversely. This suggests that a good part of the computational work could be bypassed if one knew (or had a good guess about) which columns to drop temporarily on a given iteration as non-promising. The ellipsoid based on fewer columns contains $\mathcal{E}$, and hence its use accelerates convergence. Another motivation connected with the first step of the build-up algorithm is a rule for selecting a promising set of basic columns; we believe it will speed up convergence of some practical problems, especially those with a large number of columns. Several versions of the rule lead to finite convergence, see Theorem 7.

Zikan and Cottle [13] propose the *box* method to select "likely" columns to keep in. Given an interior point y^t in $(\mathcal{D})$, they choose m columns from A corresponding to the m closest hyperplanes in $(\mathcal{D})$ to form a "box" or parallelepiped around y^t. Then the moving direction is generated by maximizing yb subject to the box constraints. This approach uses a subset of the columns and replaces the ellipsoid by a box.

Another approach, the *build-down* scheme, proposed by one of the authors, [12], is also designed to reduce the computational burden. The algorithm begins with a full set of columns. Then an eliminating criterion is applied which identifies columns

guaranteed to never be basic in any optimal solution. Via this criterion, A is eventually built-down to the optimal basis (when it is unique) and does so in polynomial time.

The main idea in this paper, in contrast to the build-down scheme, is to present a build-up interior method by selecting from A, in each major iteration, a subset of hyperplanes (columns) that the current iterate would be moving in some sense towards if the proposed move from y^t in the direction $\bar{\Delta}$ were actually made. We refer to the corresponding columns as "promising". Thus, during a major iteration we work with fewer dual constraints, hence less computational effort per major iteration, including the effort to do the rank-one update. We present an affine scaling variant of the algorithm here.

As in Dikin's method, the affine scaling variant uses ellipsoids in the suboptimization problem, but the ellipsoid is modified by replacing A with A_β, a subset of "promising" columns selected from A, which are built up during the minor cycling by blocking dual constraints. Our goal is to compare this variant with the affine scaling method of Dikin [5] (in dual form). We have also looked into the analogous variant for the related Karmarkar's projective algorithm [7] and plan to make that the subject of a subsequent paper.

The well known theoretical result is that the iterates of Dikin's algorithm (y^t, x^t) converge as $t \to \infty$ to $(\bar{y}^*, \bar{x}^*)$, the optimal dual and primal feasible solutions. Proofs of convergence can be found in [1][2][5][8][11] and in Vanderbei and Lagarias [10] under somewhat weaker assumptions. The convergence ratio $\rho^t = (\bar{y}^* b - y^{t+1} b)/(\bar{y}^* b - y^t b)$ is asymptotically bounded above by $1 - 1/\sqrt{m}$ as $t \to \infty$. A proof of this asymptotic ratio of decrease can be found in this paper and in Dantzig [3].

3. Proof of Convergence of Dikin's Algorithm

Assumptions:

(A0) $b \neq 0$, $c \neq 0$, $n > m$, and every subset of m columns from A has rank m.

(A1) An interior feasible dual solution y^1 is given.

(A2) A feasible primal solution exists.

(A3) Every feasible dual basic solution is nondegenerate.

(A4) Every primal basic solution, feasible or not, is nondegenerate.

The assumptions imply the primal feasible region is bounded, but no assumption is made about boundedness of the dual space here.

We must show first that the detailed steps of the algorithm, (2.2),...,(2.4) are legal and hence can be executed iteratively, namely:

(i) D^{-1} exists; (3.1)

(ii) $(AD^{-2}A^T)^{-1}$ exists; (3.2)

(iii) $\Delta b > 0$; (3.3)

(iv) $\bar{\Delta} = (\Delta b)^{-1/2}\Delta$ is an improving direction, i.e., $y^{t+1}b > y^t b$; (3.4)

(v) $u^{t+1} = u^t - \|Dx^t\|^{-1}(D^2x^t)^T$. (3.5)

(vi) $0 < u^{t+1} \leq 2u^t$. (3.6)

Proof. We are given inductively by (2.1) that y^t is an interior point meaning $u^t = c - y^tA > 0$. Thus, $D = \mathrm{Diag}(u^t)$ has a positive diagonal and therefore (i) D^{-1} and D^{-2} exist and have positive diagonals. The rank of A by (A0) is m and so the rank of AD^{-1} (which rescales the columns j of A by $(1/u_j^t) > 0$) is also of rank m; hence $AD^{-2}A^T$ is an m by m symmetric positive definite matrix of rank m, implying (ii), $(AD^{-2}A^T)^{-1}$ exists.

The iterates satisfy certain important relationships. First

$$Ax^t = b, \tag{4.1}$$

which is obtained by substituting the expression for Δ from (2.4) into (2.3) and multiplying by A thus:

$$Ax^t = (AD^{-2}A^T)(AD^{-2}A^T)^{-1}b = b \tag{4.2}$$

which are valid steps since D^{-1} and $(AD^{-2}A^T)^{-1}$ exist. Note that

$$Dx^t = (u_1^t x_1^t, u_2^t x_2^t, ..., u_n^t x_n^t)^T \neq 0, \quad \|Dx^t\| > 0, \tag{4.3}$$

because $u^t > 0$ and $x^t \neq 0$ (since $x = 0$ in (4.2) would imply $b = 0$ contrary to (A0)). Also note that

$$0 < \|Dx^t\|^2 = (x^t)^T D^2 x^t = (\Delta AD^{-2})D^2(D^{-2}A^T\Delta^T)$$

$$= \Delta(AD^{-2}A^T)\Delta^T = \Delta(AD^{-2}A^T)(AD^{-2}A^T)^{-1}b = \Delta b. \tag{4.4}$$

Thus, (iii) is true. This implies that it is legal to use $(\Delta b)^{-1/2}$ as a factor in (3.4) and (2.2) since $(\Delta b) > 0$. Multiplying (2.2) on the right by b and noting (4.4)

$$y^{t+1}b - y^t b = (\Delta b)^{1/2} = \|Dx^t\| > 0. \tag{4.5}$$

Thus (iv) is true; that is, $\bar{\Delta} = (\Delta b)^{-1/2}\Delta$ is a strictly improving direction.

By multiplying (2.2) on the right by A and substituting $c - u^t$ and $c - u^{t+1}$ for $y^t A$ and $y^{t+1}A$, we have from (2.1):

$$u^{t+1} = u^t - (\Delta b)^{-1/2}\Delta A \tag{4.6}$$

$$= u^t - \|Dx^t\|^{-1}\Delta A \tag{4.7}$$

$$= u^t - \|Dx^t\|^{-1}(D^2 x^t)^T \tag{4.8}$$

where (4.7) follows from (4.4); and (4.8) follows by multiplying (2.3) by D^2. Hence (v) is true.

We now verify (vi) that $0 < u^{t+1} \leq 2u^t$ by showing first $0 \leq u^{t+1} \leq 2u^t$ and next that the feasible region of $(\mathcal{E})$ is contained in the interior of the feasible region of $(\mathcal{D})$.

Let $\hat{y}$ be any point that lies in the ellipsoid $\mathcal{E}$ and let $\hat{u} = c - \hat{y}A$. An example of such a point is $\hat{y} = y^{t+1}$ and $\hat{u} = u^{t+1} = c - y^{t+1}A$. Then

$$\sum_{j=1}^{n} \frac{((\hat{y} - y^t)A_{.j})^2}{(u_j^t)^2} \le 1, \tag{5.1}$$

that is

$$\sum_{j=1}^{n} \frac{(\hat{u}_j - u_j^t)^2}{(u_j^t)^2} \le 1. \tag{5.2}$$

This implies for each j:

$$\frac{(\hat{u}_j - u_j^t)^2}{(u_j^t)^2} \le 1, \tag{5.3}$$

$$-u_j^t \le \hat{u}_j - u_j^t \le u_j^t, \tag{5.4}$$

$$0 \le \hat{u}_j \le 2u_j^t. \tag{5.5}$$

In particular, $\hat{y} = y^{t+1}$ is in $\mathcal{E}$ and therefore (5.5) implies

$$0 \le u_j^{t+1} = c_j - y^{t+1}A_{.j} \le 2u_j^t. \tag{5.6}$$

What remains to show is $u^{t+1} > 0$. Assume on the contrary, that y^{t+1} is on the boundary of the dual feasible region, then there exists a $j = s$ such that

$$0 = u_s^{t+1} = c_s - y^{t+1}A_{.s}. \tag{6.1}$$

But equality holding for (6.1) for $j = s$, implies the same for (5.3) for $j = s$ and this in turn implies that all terms of the sum (5.2) must vanish except term $j = s$. Hence $u_j^{t+1} = u_j^t$ for all j except s. Recalling (2.1) and (2.2):

$$0 = u_j^t - u_j^{t+1} = (y^{t+1} - y^t)A_{.j} = (\Delta b)^{-1/2}\Delta A_{.j} \quad j \ne s. \tag{6.2}$$

Therefore, since $\Delta b > 0$, by (3.3),

$$\Delta A_{.j} = 0 \quad \text{for all} \quad j \ne s, \tag{6.3}$$

and so from (2.3)

$$x_j^t = 0 \quad \text{for all} \quad j \neq s. \tag{6.4}$$

However choosing B to be any basic set of m indices j which includes $j = s$, we have

$$b = Ax^t = A_B x_B^t = A_{.s} x_s^t,$$

which implies that x^t is a degenerate primal basic solution, contradicting (A4). Therefore (vi) is true. Q.E.D.

This completes the proof that the detailed steps of Dikin's algorithm are legal. The property that $u^{t+1} \leq 2u^t$ is not needed for iteratively executing the algorithm but will be used in the proof of convergence which we now present in the form of several theorems and lemmas.

But first some **definitions:** A *primal solution* is any x satisfying $Ax = b$. A *dual solution* is any (y, u) satisfying $yA + u = c$. A *primal* solution is feasible if $x \geq 0$; a *dual* solution is feasible if $yA \leq c$ or $u \geq 0$. Neither primal or dual solutions need be feasible; however, all dual solutions we will be considering will be feasible ones. A partition of indices $j = 1, 2, ..., n$ into two sets consisting of m indices and the remaining $n - m$ indices will be denoted by (B, N). The set A_B of columns corresponding to B is non-singular by assumption (A0) and hence forms a basis in the space of the columns of A.

The *basic* primal and dual solutions associated with some partition (B, N) will be denoted by $\bar{x}$ and $(\bar{y}, \bar{u})$. By definition, the primal solution $\bar{x} = (\bar{x}_B, \bar{x}_N)$ associated with (B, N) is *basic*, if $\bar{x}_N = 0$. The nondegeneracy assumption (A4) also states that

$$|\bar{x}_j| > 0 \quad \text{for} \quad j \in B. \tag{7.1}$$

By definition, the dual basic solution complementary to $\bar{x}$ is $(\bar{y}, \bar{u})$ if $\bar{y}A_B = c_B$ or $u_B = 0$. The non-degeneracy assumption (A3) asserts that the dual basic solution

satisfies $|\bar{u}_j| > 0$ for $j \in N$; in particular, if the dual basic solution is *feasible*, then

$$\bar{u}_N > 0. \tag{7.2}$$

Primal-dual solutions are called *complementary* if $x_j u_j = 0$ for $j = 1, 2, ..., n$.

Theorem 1. The primal-dual iterates $\{u^t, x^t\}$ tend towards complementarity, i.e., for $j = 1, 2, ..., n$

$$u_j^t x_j^t \to 0 \quad \text{as} \quad t \to \infty. \tag{8.0}$$

Proof of Theorem 1. We note that $y^t b$ is strictly monotonically increasing by (3.4) and has a finite upper bound because by (A2) primal feasible solutions exist. Therefore from (4.5) we have

$$y^{t+1} b - y^t b = \|Dx^t\| \to 0 \quad \text{as} \quad t \to \infty. \tag{8.1}$$

where $\|Dx^t\|^2 = \sum (u_j^t x_j^t)^2$ from (4.3). Q.E.D.

Note that the proofs of convergence given in Theorems 2 and 3 that follows do not require the dual feasible region to be bounded.

Theorem 2. Given S_0 any infinite subsequence of $t = \{1, 2, ..., \infty\}$, there exists a subsequence $S_* \subset S_0$ and a partition (B, N) such that $x^t = (x_B^t, x_N^t)$ tends to the primal basic solution $(\bar{x}_B, \bar{x}_N = 0)$ and (y^t, u_B^t, u_N^t) tends to the complementary dual basic feasible solution $(\bar{y}, \bar{u}_B = 0, \bar{u}_N > 0)$ using t in S_*, i.e., as $t \to \infty$ when t is restricted to successive t in the subsequence S_*.

It is convenient to break up the proof of the theorem into three lemmas.

Lemma 2.1. Given S_0, any infinite subsequence of $t = \{1, 2, ..., \infty\}$, there exists a subsequence $S_* \subset S_0$ and a partition (B, N) such that either, for $j \in N$, $u_j^t \geq \epsilon_* > 0$ for all t in S_*, or, for $j \in B$, $u_j^t \to 0$ using t in S_*.

Proof of Lemma 2.1. Define inductively for $j = 1, 2, ..., n$ the nested set of infinite subsequences $S_0 \supset S_1 \supset S_2 ... \supset S_n$ as follows:

initialize $j := 1$; S_0 given; $\epsilon_* := \infty$;

while $j \leq n$ **do begin**

$\epsilon_j := \inf(u_j^t)$ using t in S_{j-1}; (note $\epsilon_j \geq 0$)

if $\epsilon_j > 0$ then $S_j := S_{j-1}$;

if $\epsilon_j = 0$ then $S_j :=$ any infinite subsequence of S_{j-1} such that

$u_j^t \to 0$ using t in this subsequence;

$j := j + 1$;

end while;

$S_* := S_n$; $\epsilon_* := \min(\epsilon_j$ such that $\epsilon_j > 0)$.

Let $\bar{m}$ be the number of indices j such that $u_j^t \to 0$ using t in S_*. For the remaining $n - \bar{m}$ indices j, $u_j^t \geq \epsilon_* > 0$ for all t in S_*. We have two possibilities:

Case 1: Either $\bar{m} \leq m$, in which case there exists a subset N of $n - m$ indices j such that for j in N, $u_j^t \geq \epsilon_*$ for all t in S_*; the partition (B, N) is then defined by letting B denote the remaining indices j.

Case 2: Or the alternative $\bar{m} > m$, in which case there exists a subset B of m indices such that for $j \in B$, $u_j^t \to 0$ using t in S_*. The partition (B, N) is then defined by letting N denote the remaining indices j. Q.E.D.

This completes the proof of Lemma 2.1. The two lemmas that follow establish that $\bar{m} = m$ and iterates (y^t, x^t) converge using t in S_*. The complementary basic solutions associated with (B, N) are denoted by $(\bar{x}, \bar{y})$.

Lemma 2.2. If Lemma 2.1 resulted in Case 1, then Theorem 2 is true.

Proof of Lemma 2.2. Under Case 1 there is a partition (B, N) such that for $j \in N$ that $u_j^t \geq \epsilon_* > 0$ for all t in S_*. Since $(x_j^t u_j^t) \to 0$ as $t \to \infty$ by Theorem 1 and $u_j^t \geq \epsilon_* > 0$ for $j \in N$ and all t in S_*, it follows that $x_N^t \to 0 = \bar{x}_N$ using t in S_*, and it also

follows that $x_B^t = A_B^{-1}(b - A_N x_N^t) \to A_B^{-1} b = \bar{x}_B$ using t in S_*. By the nondegeneracy assumption (A4), $|\bar{x}_j| > 0$ for $j \in B$. Again noting $(x_j^t u_j^t) \to 0$ as $t \to \infty$, it follows using t in S_* that $u_B^t \to 0 = \bar{u}_B$. Finally, we have $y^t = (c_B - u_B^t)A_B^{-1} \to c_B A_B^{-1} = \bar{y}$ using t in S_* and hence $u_N^t = c_N - y^t A_N \to c_N - \bar{y} A_N = \bar{u}_N$ using t in S_*. Moreover, since $u_N^t > 0$, $u_N^t \to \bar{u}_N \geq 0$. By the nondegeneracy assumption (A3) or (7.2), $\bar{u}_N \geq 0$ implies $\bar{u}_N > 0$.

Hence using t in S_*, $x^t \to \bar{x}$, the basic primal solution with respect to (B, N), and iterates (y^t, u^t) converge for t in S_* to the basic dual feasible solution $(\bar{y}, \bar{u}_B = 0, \bar{u}_N > 0)$. Thus Theorem 2 is proved for Case 1. Q.E.D.

Lemma 2.3. If Lemma 2.1 resulted in Case 2, then Theorem 2 is true.

Proof of Lemma 2.3. Under Case 2, there is a partition (B, N) such that, for $j \in B$, $u_B^t \to 0$ using t in S_*. As in final part of proof of Lemma 2.2, $y^t = (c_B - u_B^t)A_B^{-1} \to c_B A_B^{-1} = \bar{y}$ using t in S_*. Moreover, since $u_N^t > 0$, $u_N^t \to \bar{u}_N \geq 0$. By (A3) or (7.2), $\bar{u}_N \geq 0$ implies $\bar{u}_N > 0$. Thus $u_N^t \to \bar{u}_N > 0$ using t in S_*. But $u_N^t > 0$ and $\bar{u}_N > 0$ imply for $j \in N$ there exists an $\epsilon_* > 0$ such that $u_N^t \geq \epsilon_* > 0$ for all t in S_*. But these were the conditions we assumed in the previous lemma, hence its conclusions apply, proving Theorem 2 for Case 2 as well. Q.E.D.

Theorem 3. As $t \to \infty$, (x^t, y^t) converge to primal basic and complementary dual basic feasible solutions $(\bar{x}^*, \bar{y}^*)$.

Proof of Theorem 3. In Theorem 2, choose $S_0 = \{1, 2, ..., n\}$ and let $T_* = S_*$ be the subsequence obtained for this choice of S_0, and let $(\bar{x}^*, \bar{y}^*)$ be the basic primal solution and complementary basic dual feasible solution that (x^t, y^t) converge to using t in T_*. Denote by $(\bar{x}^p, \bar{y}^p)$ for $p = 1, 2, ..., q$ the finite set of q other basic primal and complementary basic dual feasible solutions. Let $4\delta_0$ be the shortest distance between

any two extreme points $\bar{y}$; i.e.,

$$\delta_0 = (1/4) \min \|\bar{y}^i - \bar{y}^j\|, \quad i \neq j, \tag{9.1}$$

where $i, j = *, 1, 2, ..., q$.

Given any $\delta > 0$ and any extreme points $\bar{y}$, denote by $N_\delta(\bar{y})$ the δ-neighborhood of $\bar{y}$; i.e., the ball

$$N_\delta(\bar{y}) = \{y : \|y - \bar{y}\| \leq \delta\}. \tag{9.2}$$

It is not difficult to show that the δ_0-neighborhoods of $\bar{y}^*$ and $\bar{y}^p$ for $p = 1, 2, ..., q$ have no points in common. To complete the proof we will need three lemmas.

Lemma 3.1. Given any δ, $0 < \delta < \delta_0$, the count of y^t not in the interior of any of the balls $N_\delta(\bar{y}^*)$ and $N_\delta(\bar{y}^p)$ for $p = 1, 2, ..., q$ is finite.

Proof of Lemma 3.1. An iterate y^t is said to lie outside the non-overlapping balls if $\|y^t - \bar{y}^*\| > \delta$ and $\|y^t - \bar{y}^p\| > \delta$ for $p = 1, 2, ..., q$. Let Y_0 be the subsequence of all such y^t and let S_0 be the corresponding subsequence of t. If, on the contrary, the count of t in S_0 is infinite, then by Theorem 2 there would be an infinite subsequence S_* such that $y^t \to \bar{y}$ using t in S_*, implying an infinity of t in S_0 whose y^t are in a ball $\|y - \bar{y}^p\| \leq \delta$ for some p in $\{*, 1, 2, ..., q\}$, contrary to assumption.

Lemma 3.2. If the count of y^t in the ball $N_{\delta_0}(\bar{y})$ for some extreme point $\bar{y}$ is infinite, then $y^t \to \bar{y}$ using y^t in the ball.

Proof of Lemma 3.2. Note that the set of y^t lying in the ball $N_{\delta_0}(\bar{y})$ for some extreme $\bar{y}$, but outside the smaller concentric ball $N_\delta(\bar{y})$, is a subsequence of Y_0 defined in Lemma 3.1 and therefore must also be finite whatever be $\delta < \delta_0$. It follows that if the count of y^t in $N_{\delta_0}(\bar{y})$ is infinite, then given any δ, $0 < \delta < \delta_0$, there exists a t_δ such that for all $t > t_\delta$ all y^t in $N_{\delta_0}(\bar{y})$ are also all in $N_\delta(\bar{y})$. By definition, this is what we mean when we say $y^t \to \bar{y}$ using y^t in $N_{\delta_0}(\bar{y})$.

Lemma 3.3. Either $y^t \to \bar{y}^*$ as $t \to \infty$ or there exists an infinite subsequence $T_1 = \{t_1, t_2, \dots\}$ and a successor subsequence $T_2 = \{t_1 + 1, t_2 + 1, \dots\}$ such that $y^t \to \bar{y}^*$ for $t = t_k$ in T_1 and $y^t \to \bar{y}^{p_0}$ for $t = t_k + 1$ in T_2 for some $p = p_0$.

Proof of Lemma 3.3. Assume the latter case that $y^t \to \bar{y}^*$ is not true. Generate infinite subsequences $T_1' = \{t_k\}$ and $T_2' = \{t_k + 1\}$ and $\{p_k\}$ as follows:

initialize $k := 1$; $s_1 :=$ first t such that $y^t \in N_{\delta_0}(\bar{y}^*)$;

cycle:

$t_k :=$ first $t \geq s_k$ such that $y^t \in N_{\delta_0}(\bar{y}^*)$ and $y^{t+1} \in N_{\delta_0}(\bar{y}^p)$ for some p;

$p_k := p$;

$s_k :=$ first $t > t_k$ such that $y^t \in N_{\delta_0}(\bar{y}^*)$;

$k = k + 1$;

repeat cycle;

Note that s_k always exists since $y^t \to \bar{y}^*$ using t in T_*. And t_k also exists, for else $s_k, s_k + 1, s_k + 2, \dots$ would all belong to $N_{\delta_0}(\bar{y}^*)$, implying by Lemma 3.1 that $y^t \to \bar{y}^*$ as $t \to \infty$, contrary to the contrary assumption. Therefore, since the subsequence $\{p_k\}$ is infinite and there are at most q different values that the p_k can assume, there exists a $p = p_0$ such that there is an infinite subsequence of $t \in T_1'$ whose y^t are in $N_{\delta_0}(\bar{y}^*)$ and whose successor subsequence y^{t+1} are in $N_{\delta_0}(\bar{y}^{p_0})$. Hence, there exists a subsequence $T_1 \subset T_1'$ of t_k and successor sequence $T_2 \subset T_2'$ of $t_k + 1$ such that $y^t \to \bar{y}^*$ using t in T_1 and $y^{t+1} \to \bar{y}^{p_0}$ using t in T_2.

Proof of Theorem 3 continued. Let (B, N) be the basic partition associated with $\bar{y}^*$, $\bar{u}^* = c - \bar{y}^* A$ and let $(\hat{B}, \hat{N})$ be the basic partition associated with $\bar{y}^{p_0}$, $\bar{u}^{p_0} = c - \bar{y}^{p_0} A$ where p_0 is as defined in Lemma 3.3 under the contrary assumption that y^t does not converge. Since $B \neq \hat{B}$, let $j = r$ be in B and not in $\hat{B}$. Then since

for t_k in T_1, $y^{t_k} \to \bar{y}^*$ as $k \to \infty$: .

$$u_r^{t_k} = c_r - y^{t_k} A_{.r} \to c_r - \bar{y}^* A_{.r} = \bar{u}_r^* = 0, \tag{10.1}$$

$$u_r^{t_k+1} = c_r - y^{t_k+1} A_{.r} \to c_r - \bar{y}^{p_0} A_{.r} = \bar{u}_r^{p_0} > 0, \tag{10.2}$$

where $\bar{u}_r^{p_0} > 0$ because r is non-basic with respect to $(\hat{B}, \hat{N})$ and the dual basic solution by (A3) is nondegenerate. According to (3.6), $u_r^{t+1} \leq 2u_r^t$. Since $u_r^{t_k} \to 0$, it follows that $u_r^{t_k+1} \to 0$, contrary to (10.2). Therefore the contrary hypothesis of Lemma 3.3 holds, namely (x^t, y^t) do converge to a primal basic solution $\bar{x}^*$ and the complementary dual basic feasible solution $\bar{y}^*$. Q.E.D.

Essentially the same proof of Theorem 3 was given by Todd [10]. We gave the proof above to show later where to modify it for our build-up algorithm.

Theorem 4. Dikin's algorithm converges to optimal basic primal and dual solutions.

Proof of Theorem 4. It is clear that if the primal $x^t \to \bar{x}^*$ which is feasible, then the complementary conditions (1.4) are satisfied, implying convergence in the limit to optimal solutions to the primal and dual problems. Since $Ax^t = b$, we only need to prove that $x_j^t \to x_j^* \geq 0$ for all j. At the tth iteration, we have from (4.8)

$$u_j^{t+1} = u_j^t - \frac{(u_j^t)^2 x_j^t}{\|Dx^t\|} = u_j^t\left(1 - \frac{u_j^t x_j^t}{\|Dx^t\|}\right). \tag{11.1}$$

Let us assume on the contrary for some basic index r, that $x_r^t \to \bar{x}_r^* < 0$ as $t \to \infty$. By complementarity, see Theorem 1, $u_r^t \to \bar{u}_r^* = 0$. There exists a finite $\bar{t}$ such that for all $t > \bar{t}$, $x_r^t < 0$ and therefore, since $u_r^t > 0$:

$$1 - u_j^t x_j^t / \|Dx^t\| > 1 \quad \text{for all} \quad t > \bar{t}. \tag{11.2}$$

Hence, from (11.1), for all $t > \bar{t}$

$$u_r^{t+1} > u_r^t \tag{11.3}$$

Thus, we have u_r^t strictly increasing for $t = \bar{t}+1, \bar{t}+2, \ldots, \infty$, contradicting $u_r^t \to 0$ for basic index r. This completes the proof of Theorem 4. Q.E.D.

Theorem 5. The ratio of convergence

$$\rho^t = \frac{\bar{y}^* b - y^{t+1} b}{\bar{y}^* b - y^t b} \le 1 - (\frac{1}{m} + \epsilon^t)^{1/2} \tag{12.1}$$

where $\epsilon^t \to 0$ as $t \to \infty$, i.e., ρ^t is asymptotically $\le 1 - (1/m)^{1/2}$.

Proof of Theorem 5. Let (B, N) be the partition associated with optimal $(\bar{x}^*, \bar{y}^*)$ and let $B = \{j_1, j_2, \ldots, j_m\}$. Since $\bar{y}^* = c_B A_B^{-1}$ and $y^t = (c_B - u_B^t) A_B^{-1}$:

$$\bar{y}^* b - y^t b = u_B^t A_B^{-1} b = u_B^t \bar{x}_B^*. \tag{12.2}$$

From (12.1) and (4.5),

$$\rho^t = 1 - \frac{y^{t+1} b - y^t b}{\bar{y}^* b - y^t b} = 1 - \frac{(\sum_1^n (u_j^t x_j^t)^2)^{1/2}}{u_B^t \bar{x}_B^*}. \tag{12.3}$$

We may rewrite (12.3) as:

$$(u_B^t \bar{x}_B^*)(1 - \rho^t) = \left(\sum_{j=1}^{n} (u_j^t x_j^t)^2 \right)^{1/2}$$

$$\ge \left(\sum_{i=1}^{m} (u_{j_i}^t x_{j_i}^t)^2 \right)^{1/2} \tag{12.4}$$

$$\ge \left(\sum_{i=1}^{m} (u_{j_i}^t \bar{x}_{j_i}^*)^2 + \sum_{i=1}^{m} (u_{j_i}^t \bar{x}_{j_i}^*)^2 ((x_{j_i}^t / \bar{x}_{j_i}^*)^2 - 1) \right)^{1/2}. \tag{12.5}$$

Define $\lambda_i = \lambda_i^t = u_{j_i}^t \bar{x}_{j_i}^* / u_B^t \bar{x}_B^*$. Note $\sum_1^m \lambda_i = 1$, $0 < \lambda_i < 1$, because $u_B^t \bar{x}_B^* = \sum_1^m u_{j_i}^t \bar{x}_{j_i}^*$ and $u_{j_i}^t \bar{x}_{j_i}^* > 0$. Substituting $u_{j_i}^t \bar{x}_{j_i}^* = \lambda_i (u_B^t \bar{x}_B^*)$ on the RHS of (12.5) and cancelling the common factor $u_B^t \bar{x}_B^* > 0$ from both sides:

$$(1 - \rho^t) \ge \left(\sum_{i=1}^{m} (\lambda_i)^2 + \sum_{i=1}^{m} (\lambda_i)^2 ((x_{j_i}^t / \bar{x}_{j_i}^*)^2 - 1) \right)^{1/2}. \tag{12.6}$$

$$\ge (1/m + \epsilon^t)^{1/2} \tag{12.7}$$

where the first term of (12.6), $\sum (\lambda_i)^2 \ge 1/m$ since $\sum \lambda_i = 1$ and the second term denoted ϵ^t tends to 0 because $x_{j_i}^t \to \bar{x}_{j_i}^* > 0$ and the weights $0 < \lambda_i < 1$. This completes the proof of Theorem 5. Q.E.D.

4. The Build-Up Affine Scaling Algorithm

At the start of each major iteration t, we have given an interior $y = y^t$ with $u = c - y^t A > 0$. We order the columns by some criterion such as (13.0) which measures the "promise" of a column $A_{.j}$ being in the optimal basis. The subroutine is as follows:

choose-m-promising-columns:

Reorder indices $(1, 2, ..., n)$ to $(j_1, j_2, ..., j_n)$ such that

$$u^t_{j_1} \leq u^t_{j_2} \leq .... \leq u^t_{j_n}, \tag{13.0}$$

$$\beta^t := (j_1, j_2, ..., j_m) \tag{13.1}$$

Later we will show that this rule and some other alternate rules for choosing the starting $\beta = \beta^t$ all lead to finite convergence, see Theorem 7. In practical applications, the assumption that every set of m columns is independent is usually not correct and (13.1) is modified to read: choose β as the indices of the first m most promising *independent* columns. β stands for "build up". During a major iteration the initial set of indices (13.1) is built up to include additional indices. When β is used as a subscript, it means that the indices j are restricted to β.

The second subroutine is

solve-starting-basis-(β):

Here β is the set of indices of the m most promising columns. A_β at the start of a major iteration is nonsingular and therefore may serve as a basis. The detailed steps are:

$$x_\beta = (A_\beta)^{-1} b; \quad \pi = c_\beta (A_\beta)^{-1}; \tag{14.0}$$

$$r = \arg\min_{j_i} (x_\beta). \tag{14.1}$$

The third subroutine is

solve-ellipsoid-(y^t, β):

Here β refers to the initial set of promising indices plus those built-up so far during the minor cycles within a major iteration. The ellipsoid subproblem to solve is

$$(\mathcal{E}_\beta) \qquad \text{maximize} \quad yb \quad \text{subject to} \quad y \in \mathcal{E}_\beta = \{y : \|(y - y^t)A_\beta D_\beta^{-1}\| \leq 1\}.$$

The subroutine steps to solve $(\mathcal{E}_\beta)$ are:

$$D_\beta = \text{Diag}(u_\beta^t); \tag{15.1}$$

$$\Delta = b^T (A_\beta D_\beta^{-2} A_\beta^T)^{-1}; \tag{15.2}$$

$$\bar{\Delta} = (\Delta b)^{-1/2} \Delta. \tag{15.3}$$

The fourth subroutine is

find-blocking-constraint-(s):

At the start of a major iteration, *ratio* is set equal to $+\infty$ and the subroutine scans from y^t in the direction $\bar{\Delta}$ looking for the first blocking constraint $j = s$; at the start of each minor cycle, *ratio* is set equal to 1 and the subroutine then scans the line segment y^t to $y^t + \bar{\Delta}$ looking for the first blocking constraint $j = s$. If no blocking constraint is encountered, $s = 0$. The detailed searching steps are:

> $s := 0$;
>
> $v := \bar{\Delta} A$;
>
> **for** $j := 1$ **to** n **do begin**
>
>   **if** $((j \notin \beta)$ **and** $(v_j > 0)$ **and** $(u_j^t / v_j \leq ratio)$ **then begin**
>
> $ratio := u_j^t / v_j; \quad s := j;$
>
>   **end**;
>
> **end for**.

The Main Program of the build-up affine scaling method is as follows:

Program;

Input: dual feasible solution y^1 and $u^1 = (c - y^1 A)^T > 0$;

Initialize $t := 1$ and $opt := false$;

While $((opt = false)$ **do begin**

choose-m-promising-columns;

solve-starting-basis-(β) for β and r;

if $((x_{j_r} \geq 0)$ and $(c - \pi A \geq 0))$ **then begin**

output ("optimum", x_β, $x_N = 0, \pi$); $opt := true$;

end;

if $(opt = false)$ **then begin**

$\bar{\Delta} := (A_\beta)^{-1}_{r.}$;

$ratio := +\infty$;

find-blocking-constraint-(s);

if $s = 0$ **then begin**

output ("error, primal infeasible"); $opt := true$;

end;

while $s > 0$ **do begin**

$\beta := \beta \cup \{s\}$;

solve-ellipsoid-(y^t, β) for $\bar{\Delta}$;

$ratio := 1$;

find-blocking-constraint-(s);

end (end of a minor cycle);

$x^t_j := (\Delta A_{.j})/(u^t_j)^2$ for $j \in \beta$;

$x^t_j := 0$ for $j \notin \beta$;

$y^t := y^t + \bar{\Delta}$; $u^t := c - y^t A$;

$t := t + 1$;

end if;

end while (end of a major iteration);

end program.

5. Proof of Finite Convergence of Build-Up Algorithm

Theorem 6. Independent of the rule used to *choose-m-promising-columns*, the iterative process either terminates in a finite number of iterations with the optimal primal and dual solutions or converges to them in the limit.

Proof of Theorem 6. Let us assume it does not converge in a finite number of iterations. Each major iteration t, after it checks and finds that the starting basis is not optimal, engages in a number of minor cycles building up β^t (the indices of the constraints used to define the ellipsoid $\mathcal{E}_\beta$) until there is no constraint blocking movement from y^t to $y^t + \bar{\Delta}$, the optimal point of $(\mathcal{E}_\beta)$: i.e., $y^t + \bar{\Delta}$ is an interior point of $(\mathcal{D})$. The count of minor cycles within a major iteration cannot be more than $n - m$ because this is the number of remaining constraints $yA_{.j} \le c_j$ whose index j might be used to augment starting β. Therefore minor cycling always terminates with a move from y^t to y^{t+1} where the formulae for this move are the analogs of (2.1) to (2.4) of Dikin's algorithm.

$$\beta = \beta^t, \quad D_\beta = \text{Diag}(u^t_\beta), \tag{16.0}$$

$$\Delta = b^T (A_\beta D_\beta^{-2} A_\beta^T)^{-1}, \tag{16.1}$$

$$\bar{\Delta} = (\Delta b)^{-1/2} \Delta, \tag{16.2}$$

$$y^{t+1} = y^t + \bar{\Delta}, \tag{16.3}$$

$$x^t_\beta = D_\beta^{-2} A_\beta^T \Delta^T; \quad x^t_j = 0 \quad \text{for} \quad j \notin \beta. \tag{16.4}$$

In addition, we have the analogs to (4.1) to (4.4)

$$b = A_\beta x^t_\beta = Ax^t, \tag{17.0}$$

$$0 < \|D_\beta x^t_\beta\|^2 = \Delta b, \quad \|D_\beta x^t_\beta\| = \|Dx^t\|, \tag{17.1}$$

because $x_j^t = 0$ for $j \notin \beta$. The proofs that D_β^{-1} and the $(A_\beta D^{-2} A_\beta^T)^{-1}$ exist are the same as those without the subscript β, see (3.1) and (3.2) implying that steps (16.1) and (16.2) are legal. The proofs that

$$\Delta b > 0, \tag{17.2}$$

$$\bar{\Delta} = (\Delta b)^{-1/2} \Delta \quad \text{is an improving direction,} \tag{17.3}$$

$$y^{t+1} b - y^t b = (\Delta b)^{1/2} = \|D_\beta x_\beta^t\| = \|D x^t\| > 0 \tag{17.4}$$

follow the proofs of (3.3) and (3.4). The analogs to (3.5) and (3.6) are

$$u_\beta^{t+1} = u_\beta^t - \|D_\beta x_\beta^t\|^{-1} (D_\beta^2 x_\beta^t)^T, \tag{17.5}$$

$$0 < u_\beta^{t+1} \le 2u_\beta^t, \quad 0 < u^{t+1}, \tag{17.6}$$

where $u_j^{t+1} > 0$ for $j \notin \beta$ is true because the minor cycles do not terminate until there is no blocking constraint, i.e., $u_j^{t+1} = c_j - y^{t+1} A_{.j} > 0$ for all $j \notin \beta$.

Under the assumption of no finite termination, we now show that Theorems 1 to 5 of Dikin's algorithm are true for the build-up algorithm. However the reasons *why* they are true will need more explanation in places and this we will now proceed to do.

That $u_j^t x_j^t \to 0$ as $t \to \infty$, Theorem 1, follows from $(y^{t+1} b - y^t b) \to 0$ and (17.4). The proof of Theorem 2 is the same. It states that given any infinite subsequence S_0 of t, there exists a subsequence $S_* \subset S_0$ such that (x^t, y^t) converge using t in S_* to $(\bar{x}, \bar{y})$, a pair of basic primal and complementary basic dual feasible solutions.

Next to show Theorem 3 for the build-up algorithm that (x^t, y^t) converges as $t \to \infty$ to a pair of basic primal and complementary basic dual feasible solutions, we note all steps of the proof apply up to the very last step where relation (3.6) that $u_r^{t+1} \le 2u_r^t$ for $t = t_k$ in the subsequence T_1 is used to obtain a contradiction. However the analog of (3.6) for the build-up algorithm is (17.6) which states $u_r^{t+1} \le 2u_r^t$

provided $r \in \beta = \beta^t$. We claim that except for a finite subsequence of t in T_1 that r is indeed in β^t; if so then this is clearly sufficient to complete the proof of Theorem 3. Suppose on the contrary there is an infinite subsequence R_1 in T_1 such that , for t in R_1, r not in β^t. Since $R_1 \subset T_1$ is infinite, its convergence properties are the same as T_1 and therefore $x_r^t \to \bar{x}_r^*$ and $|\bar{x}_r^*| \neq 0$ because r is basic index in (B, N). On the other hand, $x_r^t \to \bar{x}_r^* = 0$ for t in R_1 because the algorithm sets all $x_j^t = 0$ for $j \notin \beta^t$ which contradicts $|\bar{x}_r^*| \neq 0$ for $r \in B$.

Also requiring explanation is Theorem 4 which asserts that (x^t, y^t) converge to complementary optimal basic primal and dual solutions. The proof assumes, on the contrary, $\bar{x}^*$ is not feasible and that there is some basic index $j = r$ such that $x_r^t \to \bar{x}_r^* < 0$ as $t \to \infty$. To get a contradiction, relation (11.1) was used; its analog (17.5) is applicable provided $r \in \beta^t$. However there can be at most a finite count of t such that basic $r \notin \beta^t$ and a corresponding count of $x_r^t = 0$, because if the set S of $r \notin \beta^t$ were infinite, this would imply $x_r^t \to \bar{x}_r^* = 0$, but this limit is the same as that for all t, contradicting $|\bar{x}_r^*| \neq 0$. Therefore there exists a finite t_0 such that for all $t > t_0$ that $B \in \beta^t$ and (17.6) is true for all $r \in B$, and the rest of the proof of the Theorem 4 can now be applied without further change.

Finally the proof of Theorem 5 about the asymptotic ratio of convergence being $\rho^t \leq 1 - (1/m)^{1/2}$ may be used without change. Q.E.D.

This completes the proof of Theorem 6 which asserts that if the build-up algorithm does not detect an optimal basic feasible solution at start of some major iteration t and terminate, then it will converge in the limit to such a solution.

We now present several alternative rules that rank the promise of column j being in the optimal basis according to the ratio $(x_j^t)^\mu/(u_j^t)^\nu$ from high to low:

Rule 1 $(\mu, \nu) = (0, 1)$:

$$(1/u_{j_1}^t) \geq (1/u_{j_2}^t) \geq \ldots \geq (1/u_{j_n}^t).$$

Rule 2 $(\mu, \nu) = (1, 0)$:

$$(x_{j_1}^{t-1}/1) \geq (x_{j_2}^{t-1}/1) \geq \dots \geq (x_{j_m}^{t-1}/1).$$

Rule 3 $(\mu, \nu) = (1, 1)$:

$$(x_{j_1}^{t-1}/u_{j_1}^{t}) \geq (x_{j_2}^{t-1}/u_{j_2}^{t}) \geq \dots \geq (x_{j_m}^{t-1}/u_{j_m}^{t}).$$

Theorem 7. If the starting $\beta^t = (j_1, j_2, \dots, j_m)$ is chosen by any of the above rules, the build-up algorithm will terminate in a finite number of iterations with the optimal primal and dual solutions.

Proof of Theorem 7. Let the optimal partition be (B, N). Assume on the contrary that the particular rule chosen does not lead to finite termination, then by Theorem 6 as $t \to \infty$, (x_B^t, x_N^t) converge to optimal $(\bar{x}_B^* > 0, \bar{x}_N^* = 0)$; and (u_B^t, u_N^t) converge to $(\bar{u}_B^* = 0, \bar{u}_N^* > 0)$. Let

$$\epsilon_0 = (1/2) \min(\bar{x}_{i \in B}^*, \bar{u}_{j \in N}^*).$$

There exists a t_0 such that for all $t > t_0$:

$$u_j^t < \epsilon_0 \quad j \in B \quad \text{and} \quad u_j^t > \epsilon_0 \quad j \in N.$$

$$x_j^t > \epsilon_0 \quad j \in B \quad \text{and} \quad x_j^t < \epsilon_0 \quad j \in N.$$

$$x_j^t/u_j^t > 1 \quad j \in B \quad \text{and} \quad x_j^t/u_j^t < 1 \quad j \in N.$$

Therefore, contrary to the assumption of convergence in an infinite number of iterations, we see that under any one of the three rules, the basis associated with the starting β^t for all $t > t_0$ is optimal and the algorithm will detect this fact by iteration $t \leq t_0$. Q.E.D.

6. Number of Minor Cycles

The count $h^t \geq m$, the number of indices in $\beta = \beta^t$ build up by the end of major iteration t, is not fixed and can vary up and down from one iteration to the next. We expect in early iterations that h^t will be somewhat large, but as t increases h^t will likely be close to m. Also, one may allow several columns to enter A_β simultaneously, instead of one at a time, and allow some columns deleted from A_β in order to keep the size of A_β under control. In the following, we show, at least heuristically, that the size of A_β might be controlled to no more than $2m$ for problems with $n > 2m$. This is based on the following theorem:

Theorem 8. Let $\Omega = \{y \in R^m : c - yA \geq 0\}$ where $A \in R^{m\times n}$ and $n = 2m + 1$, and let y^0 be any interior point in Ω. Also, let A_β denote an $m \times (2m)$ submatrix of A. Then, there exists a particular A_β such that

$$\mathcal{E}_\beta = \{y : \|(y - y^0)A_\beta D_\beta^{-1}\| \leq 1\} \subset \Omega$$

where $D_\beta = \mathrm{Diag}(c_\beta - y^0 A_\beta)$.

Proof. We know that

$$\mathcal{E}_\beta \subset \{y : c_\beta - yA_\beta \geq 0\}.$$

Therefore, we only need to show that

$$\mathcal{E}_\beta \subset \{y : c_s - yA_{.s} \geq 0\} \quad \text{for} \quad s \notin \beta.$$

Let $a_j = A_{.j}$, $D = \mathrm{Diag}(c - y^0 A)$ and let σ be the vector whose j-th component is

$$\sigma_j = a_j^T (AD^{-2}A^T)^{-1} a_j / D_{jj}^2 > 0,$$

which are diagonal components of the projection matrix

$$D^{-1}A^T(AD^{-2}A^T)^{-1}AD^{-1}.$$

The trace of the projection matrix satisfies

$$\sum_{j=1}^{2m+1} \sigma_j = m.$$

Therefore, there exists at least one $j = s$ such that $\sigma_s < 1/2$. Define β to be the $2m$ indices which excludes s. We show $\mathcal{E}_\beta \subset \{y : c_s - ya_s > 0\}$, i.e., constraint s does not block $\mathcal{E}_\beta$ in any direction. As shown in Reference [12], it suffices to prove that

$$\bar{\sigma}_s = a_s(A_\beta D_\beta^{-2} A_\beta^T)^{-1} a_s / D_{ss}^2 < 1.$$

This is because the following argument. The maximal value of the problem:

$$\text{maximize} \quad (y - y^0)a_s \quad \text{subject to} \quad y \in \mathcal{E}_\beta$$

is exactly

$$(\bar{y} - y^0)a_s = \sqrt{a_s(A_\beta D_\beta^{-2} A_\beta^T)^{-1} a_s}$$

where $\bar{y}$ denotes the maximum. Thus, if

$$c_s - y^0 a_s = D_{ss} > \sqrt{a_s(A_\beta D_\beta^{-2} A_\beta^T)^{-1} a_s} = (\bar{y} - y^0)a_s,$$

then

$$c_s > \bar{y}a_s \geq ya_s \quad \text{for all} \quad y \in \mathcal{E}_\beta.$$

Note that

$$(AD^{-2}A^T)^{-1} = (A_\beta D_\beta^{-2} A_\beta^T + a_s a_s^T / D_{ss}^2)^{-1}$$

$$= (A_\beta D_\beta^{-2} A_\beta^T)^{-1} - (A_\beta D_\beta^{-2} A_\beta^T)^{-1} a_s a_s^T (A_\beta D_\beta^{-2} A_\beta^T)^{-1} / (D_{ss}^2(1 + \bar{\sigma}_s)). \qquad (18.1)$$

Multiplying (18.1) by a_s^T and a_s from the left and right, and then dividing it by D_{ss}^2, we have

$$\sigma_s = \bar{\sigma}_s - \bar{\sigma}_s^2 / (1 + \bar{\sigma}_s)$$

which yields

$$\bar{\sigma}_s = \sigma_s/(1-\sigma_s).$$

Thus, $\sigma_s < 1/2$ indicates that

$$\bar{\sigma}_s < 1$$

This completes the proof of Theorem 8. Q.E.D.

Theorem 8 states that the ellipsoid $\mathcal{E}_\beta$ inscribed in polytope Ω with $2m+1$ constraints can be constructed using at most $2m$ constraints, which suggests but does not establish that h^t will be at most $2m$. This helps when faced by an $n > 2m$ LP problem because if the number of components of β becomes greater than $2m$ the calculation of $A_\beta D_\beta^{-2} A_\beta^T$ will be a big burden as it is in conventional interior-point algorithms. Note that the build-up method can also be used for optimization problems where linear constraints are generated during the course of the algorithm.

A rank-one updating method should be used to efficiently refactorize

$$(A_\beta D_\beta^{-2} A_\beta^T)^{-1}$$

and recalculate Δ after A_β is augmented by a blocking dual constraint s. This efficiency can be achieved using any one of several techniques (see, for example, Gill et al. [6] and Ng [9]). Also, at the beginning of each major iteration we need to invert a basis typically in factorized form. If the basis differs from the one at the beginning of the previous major iteration by only few columns, we can apply a few rank-one updates to obtain its inverse in factorized form as well.

References

[1] I. Adler, N. Karmarkar, M. G. C. Resende and G. Veiga, "An implementation of Karmarkar's algorithm for linear programming," Working Paper, Operations Research Center, University of California (Berkeley, CA, 1986).

[2] E. R. Barnes, "A variation on Karmarkar's algorithm for linear programming," *Mathematical Programming* **36** (1986) 174-182.

[3] G. B. Dantzig, "Dikin's interior method for solving LP," manuscript, Department of Operations Research, Stanford University (Stanford, CA, 1988).

[4] G. B. Dantzig, *Linear Programming and Extensions* (Princeton University Press, Princeton, NJ, 1963).

[5] I. I. Dikin, "Iterative solution of problems of linear and quadratic programming," *Doklady Akad. Nauk USSR* **174** (1967), Translated in *Soviet Math. Doklady* **8** (1967) 674-675.

[6] P. E. Gill, W. Murray, M. A. Saunders, and M. H. Wright, "Maintaining LU factors of a general sparse matrix," *Linear Algebra and its Applications* **88/99** (1987) 239-270.

[7] N. Karmarkar, "A new polynomial-time algorithm for linear programming," *Combinatorica* **4** (1984) 373-395.

[8] C. L. Monma and A. J. Morton, "Computational experimental with a dual affine variant of Karmarkar's method for linear programming," Technical Report, Bell Communications Research (Morristown, NJ, 1987).

[9] E. Ng, "On the solution of sparse linear least-squares problems," Presentation at Stanford, Mathematical Science Section, Oak Ridge National Labs (Oak Ridge, TN, 1988).

[10] R. J. Vanderbei and J. C. Lagarias, "I. I. Dikin's convergence result for the affine-scaling algorithm," manuscript, AT&T Bell Laboratories (Murray Hill, NJ, 1988).

[11] R. J. Vanderbei, M. S. Meketon and B. A. Freedman, "A modification of Karmarkar's linear programming algorithm," *Algorithmica* 1 (1986) 395-407.

[12] Y. Ye, "A 'build-down' scheme for linear programming," manuscript, Department of Management Sciences, The University of Iowa (Iowa City, IA, 1988), to appear in *Mathematical Programming*.

[13] K. Zikan and R. W. Cottle, "The box method for linear programming: Part I—basic theory," Technical Report SOL 87-6, Department of Operations Research, Stanford University (Stanford, CA, 1987).

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER: SOL 90-4	2. GOVT ACCESSION NO.
3. RECIPIENT'S CATALOG NUMBER	
4. TITLE (and Subtitle): A Build-Up Interior Method for Linear Programming: Affine Scaling Form	5. TYPE OF REPORT & PERIOD COVERED: Technical Report
	6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s): George B. Dantzig and Yinyu Ye	8. CONTRACT OR GRANT NUMBER(s): N00014-89-J-1659
9. PERFORMING ORGANIZATION NAME AND ADDRESS: Department of Operations Research - SOL, Stanford University, Stanford, CA 94305-4022	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS: 1111MA
11. CONTROLLING OFFICE NAME AND ADDRESS: Office of Naval Research - Dept. of the Navy, 800 N. Quincy Street, Arlington, VA 22217	12. REPORT DATE: February 1990
	13. NUMBER OF PAGES: 29 pp.
	15. SECURITY CLASS. (of this report): UNCLASSIFIED
	15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

This document has been approved for public release and sale; its distribution is unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

linear programming; interior method; affine scaling.

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

(please see reverse side)

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

A BUILD-UP INTERIOR METHOD FOR LINEAR PROGRAMMING: AFFINE SCALING FORM

George B. Dantzig and Yinyu Ye

February 1990

Abstract

We propose a *build-up* interior method for solving an m equation n variable linear program which has the same convergence properties as their well known analogues in dual affine and projective forms but requires less computational effort. The algorithm has three forms, an *affine scaling* form, a *projective scaling* form, and an *exact* form (that uses pivot steps). In this paper, we present the first of these. It differs from Dikin's algorithm of dual affine form in that the ellipsoid chosen to generate the improving direction $\bar{\Delta}$ in dual space is constructed from only a subset of the dual constraints.

At the start of each major iteration t, we are given an interior iterate y^t. A selection of m dual constraints is made using an "order-columns" rule as to which constraints show "the most promise" of being tight in the optimal dual solution. An ellipsoid centered at y^t is then inscribed in convex region defined by these promising constraints and an improving direction $\bar{\Delta}$ computed that points to the optimal point $y^t + \bar{\Delta}$ on the ellipsoid boundary. Minor cycling within a major iteration is then started.

During a minor cycle, the constraints selected to define the ellipsoid centered at y^t is built up to include the constraint (whenever there is one) that first blocks feasible movement from y^t to $y^t + \bar{\Delta}$. If one blocks, it is used to augment the set of promising constraints and the ellipsoid is revised; the improving direction $\bar{\Delta}$ is recomputed by means of a rank-one update, and the minor cycle repeated until none blocks movement from y^t to $y^t + \bar{\Delta}$. When none blocks, the minor cycling ends. $y^{t+1} = y^t + \bar{\Delta}$ initiates the next major iteration. Major iterations stop when an optimum solution is reached. We prove this will occur in a finite number of iterations.